

ABSTRACT OF THE INVENTION

RETRIEVAL OF RECORDS USING PHRASE CHUNKING

5 Methods are provided for generating phrase chunking rules for titles of records in a database. According to one method, the title of each record in a first set of records is part-of-speech tagged, and a plurality of phrase chunking rules are created based on patterns of part-of-speech tags in the tagged titles. The phrase chunking rules are applied to the titles of records in a second set of records so as to generate indexes for the records in the second set
10 10 of records. In a preferred embodiment, the phrase chunking rules are modified if coverage of the second set of records by the phrase chunking rules does not reach a predetermined threshold. Also provided are methods for retrieving records from a database and systems for generating phrase chunking rules.